Toxicogenomics

Liu, Zhichao; Huang, Ruili; Roberts, Ruth; Tong, Weida

[Link to publication on Research at Birmingham portal](#)

CellPress

Opinion

Toxicogenomics: A 2020 Vision

Zhichao Liu,[1,*] Ruili Huang,[2] Ruth Roberts,[3,4] and Weida Tong[1,*]

Toxicogenomics (TGx) has contributed significantly to toxicology and now has great potential to support moves towards animal-free approaches in regulatory decision making. Here, we discuss *in vitro* TGx systems and their potential impact on risk assessment. We raise awareness of the rapid advancement of genomics technologies, which generates novel genomics features essential for enhanced risk assessment. We specifically emphasize the importance of reproducibility in utilizing TGx in the regulatory setting. We also highlight the role of machine learning (particularly deep learning) in developing TGx-based predictive models. Lastly, we touch on the topics of how TGx approaches could facilitate adverse outcome pathways (AOP) development and enhance read-across strategies to further regulatory application. Finally, we summarize current efforts to develop TGx for risk assessment and set out remaining challenges.

Toxicogenomics in Regulatory Application: Challenges and Opportunities

Animal models are used to assess and avoid risk to humans from exposure to potential hazards, but their use is under constant review, especially in the light of some reports of poor extrapolation for complex endpoints, such as hepatotoxicity and carcinogenicity. Consequently, 21st century toxicology emphasizes alternative means of risk assessment and the promotion of the 3Rs (replacement, reduction, and refinement of animals in toxicology testing) [1]. In Europe, great efforts have been made to advance the 3Rs with the aim of developing animal-free risk assessment methodologies. To this end, several high-profile programs are underway, such as the Framework Programme 7 (FP7), Horizon 2020, and some public–private partnerships, including Safety Evaluation Ultimately Replacing Animal Testing (SEURAT-1) and the Innovative Medicines Initiative (IMI). Furthermore, a series of EU Legislative directives have been developed and implemented over the past three decades, with an emphasis on moving away from animal testing; since 2013, animal models have been prohibited for testing cosmetics or household products in the EU, as well as in Israel and India [2]. In the US, government-initiated efforts comprise advanced regulatory sciences proposed by the US FDA [3] and Tox21 [4] [which involves four government agencies, including the Environmental Protection Agency (EPA), National Center for Advancing Translational Sciences, National Institute of Environmental Health Sciences, and the FDA] and ToxCast [5] (by the EPA). These ongoing efforts actively advocate and promote *in silico* and *in vitro* approaches, including **toxicogenomics (TGx)** (see Glossary), for prioritization and also for a potential application in risk assessment.

TGx, as a subdiscipline of toxicology, has been successfully implemented to address critical issues and questions in a broad spectrum of toxicology. The rapid advancement of next-generation sequencing (NGS) technologies has gained traction in clinical application, particularly in personalized cancer diagnosis and prognosis, offering great opportunities for precision medicine. Meanwhile, the entire toxicology field has also been impacted by the rapid development with these advanced sequencing technologies. Equipped with innovative technologies in

Highlights

Together with the promotion of non-animal testing, *in vitro* toxicogenomics (TGx) may play a vital role in the next-generation risk assessment paradigm.

A strategic shift in risk assessment provides an unprecedented opportunity for repositioning TGx in the regulatory setting.

As the emerging technique continues to impact the TGx field, novel genomic features such as miRNAs, ncRNAs, and circular RNAs may provide more resolution towards better understanding of the underlying mechanisms of toxicological processes.

Advances in machine learning and artificial intelligence are gaining ground for their applicability in biomedical fields. In the near future, these advances may be further applied in the TGx field to improve predictive power.

[1]National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas 72079, USA
[2]National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, Maryland 20850, USA
[3]ApconiX, BioHub at Alderley Park, Alderley Edge, SK10 4TG, UK
[4]University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

*Correspondence:
zhichao.liu@fda.hhs.gov (Z. Liu) and
weida.tong@fda.hhs.gov (W. Tong).

CrossMark

both genomics and high-throughput methodologies, TGx provides an unprecedented opportunity for improved risk assessment and regulatory decision making. Moving towards the concept of an animal-free, next-generation risk assessment paradigm, it is of great importance to rethink and reposition TGx to meet regulatory needs.

Figure 1 summarizes some challenges and opportunities in moving from considering available technologies to the use of TGx in the regulatory decision-making process. While the potential uses of TGx results in the regulatory environment is multifaceted, current efforts and progress are mainly in the area of mechanistic and predictive information. Mechanistic information can be presented in a variety of forms, such as mode-of-action, critical toxicity related pathways, and gene functions. Among these, advanced **adverse outcome pathways (AOPs)** hold the promise to elucidate underlying toxicity mechanisms, enabling screening assay development [6]. On the predictive side, similarly, there exists a broad range of machine learning methods. Notably, deep learning (part of the broader family of machine learning methods based on learning data representations as opposed to task-specific algorithms) could be a promising approach to advance TGx. Irrespective of the approach being used and the specific application intended, reliable and consistent reproduction of results is the prerequisite for successful application of TGx results in the regulatory setting. Thus, it is imperative to establish and implement a reliable and reproducible protocol for *in vitro* TGx studies. With the rapid advancement and evolution of genomic technologies, there are significant challenges in establishing consistency and transferability in approaches between fields and between labs and users. In the following sections, we will discuss some of the options and opportunities in these fields.

In Vitro TGx

As highlighted earlier, risk assessment is moving away from animal testing to alternative methodologies. Notably, *in vitro* assay systems derived from animal or human tissues are proving increasingly useful as an adjunct within the longer-term animal-based testing paradigm. However, these *in vitro* approaches are challenged due to uncertainty and limited transferability power to *in vivo* models [7]. Table 1 summarizes some representative cell-based *in vitro* models used as a starting point for the generation of TGx data. Notably, *in vitro* TGx approaches are still dominated by liver models; in the absence of a comprehensive, head-to-head comparison it remains difficult to conclude whether liver models are indeed the most informative. Also, the
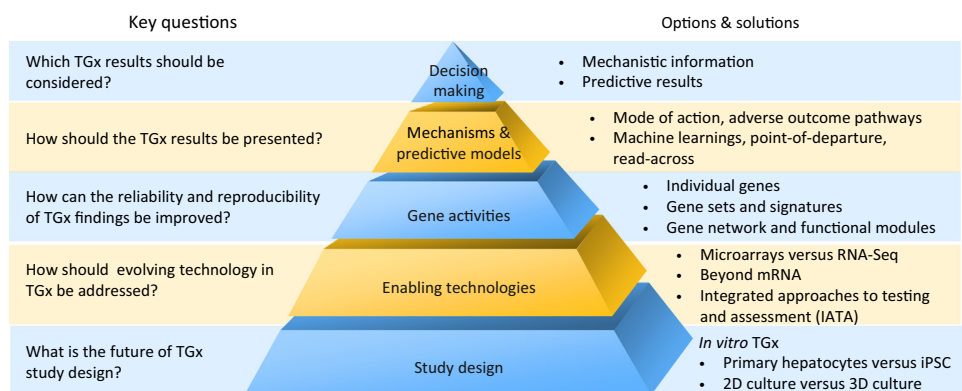
Trends in Pharmacological Sciences

Figure 1. The Pyramid of Toxicogenomics (TGx) towards Regulatory Decision Making. Some outstanding questions and potential solutions for promoting TGx in respect of decision making are shown. iPSC, Induced pluripotent stem cell.

Glossary

Adverse outcome pathways (AOPs): a conceptual framework connecting a molecular initiating event and key events with outcome and adverse effects in risk assessment.

DrugMatrix: one of the largest toxicogenomic reference resources, consisting of 638 compounds tested under microarray technology and their corresponding pathology data in rat, which covers 137 mechanism of toxicity-related pathways and 50 pathological endpoints.

FAIRsharing community: a web-based, searchable portal of three interlinked registries, containing both in-house and crowdsourced manually curated descriptions of standards, databases, and data policies, combined with an integrated view across all three types of resource.

Gene Expression Omnibus (GEO): the world's largest functional genomics data repository developed by the National Center for Biotechnology Information.

Idiosyncratic toxicity: is not dose-dependent and unpredictable. Idiosyncratic toxicity is caused by drug- and patient-related risk factors. Drug-related risk factors include metabolism, bioactivation and covalent binding, and the inhibition of key cell functions. Patient-related risk factors include genetic background, underlying disease, age, gender, comedications, and activation of the innate immune system.

***In vitro* to *in vivo* extrapolation (IVIVE):** can be broadly defined as an approach extrapolating the experimental results or observations made *in vitro* to predict *in vivo* phenomena qualitatively or quantitatively.

KEGG: Kyoto Encyclopedia of Genes and Genomes database, a collection of data resources that takes account of the complex relationship among biological pathways, diseases, and chemical substances/drugs.

OmicsMapNet: an approach to transform omics data to 2D images as an input for building deep convolutional neural network (CNN).

Open TG-GATEs: a large-scale toxicogenomics database that stores gene expression profiles, pathological

quality and utility of the different types of cells used are all influenced by multiple factors, such as chemical properties, cell differentiation, tissue variation, measured toxicity endpoints, and experimental protocols.

Early efforts on TGx testing system comparisons have demonstrated the complexity and challenges imposed by the variability among cell types, gene profiling platforms, and measured endpoints [8]. Often, conclusions varied due to the limited number of compounds studied and the different types of cells and cultures. Some recent comparisons among different TGx testing systems were conducted with large-scale TGx studies, including multiple cell cultures [9,10]. For example, Sutherland *et al.* [9] comprehensively assessed the concordance of drug-induced transcriptional response in the rat, mouse liver, and cultured hepatocytes with **Gene Expression Omnibus (GEO)**, **Open TG-GATEs** [11][iii] and **DrugMatrix**. The authors suggested that better concordance could be obtained at the coexpression network level of RNA than at individual genes and gene-set levels. Nonetheless, most TGx studies have concluded that the variability is mainly dependent on measured endpoints [12], implying that the selection of *in vitro* TGx model system should largely be driven by the type of perturbations being studied.

It is valuable to highlight that some novel *in vitro* models, such as 3D cultures, stem cells, organoids, and physiologically based pharmacokinetic modeling have great potential to improve the stability of *in vitro* assay systems (Table 1) [13,14]. Unfortunately, the evaluation of the use of these novel models in TGx is limited by the number of compounds tested so far. Although the future of these types of *in vitro* TGx are promising and could play a critical role in regulatory decision making, comprehensive and comparative analyses with study designs across various novel models need to be conducted to warrant their utility and define the applications where they are fit-for-purpose [15].

Technology Advancements and Their Impact on TGx

Technology advancements make TGx approaches increasingly powerful and cost effective. The interplay between technology evolution and toxicology application continues to drive TGx into high-throughput applications for mechanistic investigation and multigene biomarker discovery. There is a steep learning curve between novel technology to real-world application; the key challenge to overcome in order to take full advantage of these novel technologies is in choosing fit-for-purpose approaches in balancing cost and effect.

Microarray versus RNA-Seq

Rapid technological advancement, along with a sharp decline in cost, has placed NGS at the center stage of genomics research in the broad biomedical field [16]. This raises the question as to whether the TGx field is ready to migrate from microarray to NGS-based approaches. Currently, the answer remains ambiguous and complex. On the one hand, microarray technology has a proven track record spanning the past two decades with a vast amount of accumulated microarray data that is publicly available. On the other hand, **RNA-Seq** has demonstrated advantages in detecting low-expressed genes compared with microarrays [12]. So, can we 'bridge' by applying biomarkers generated from microarray platforms to data derived from RNA-Seq? Since the technology is rapidly evolving, this bridging challenge will be repeated in each iteration. Thus, it would be valuable to focus on understanding the potential transferability of biomarkers from old technologies to new ones and vice versa. Via this route, the predictive sustainability across evolved technologies can be understood and the rich data sets accumulated from past investment can be leveraged.

data, and biochemical testing results derived from rat *in vivo* and *in vitro* (primary rat hepatocytes, primary human hepatocytes) exposure to 170 compounds at multiple time and dosages points.

Phase II enzymes: phase II enzymes oversee conjugation reactions, which generally serve as a detoxifying step in drug metabolism. Phase II enzymes are mainly transferases.

Registration, Evaluation, Authorization and Restriction of Chemicals (REACH): a European Union regulation that aims to address the production and use of chemical substances, and their potential impacts on both human health and the environment.

RNA-Seq: uses next-generation sequencing (NGS) to reveal the presence and quantity of RNA in a biological sample at a given moment, which is used to analyze the continuously changing cellular transcriptome.

Toxicogenomics (TGx): a subdiscipline of pharmacology that applies genomic technologies (e.g., genetics, genome sequence analysis, gene expression profiling, proteomics, metabolomics, and related approaches) to study the adverse effects of environmental and pharmaceutical chemicals on human health and the environment.

Treemap approach: an information visualization method for displaying hierarchical data using nested figures such as rectangles.

Xenobiotic metabolism: xenobiotic metabolism is in charge of removing xenobiotics from the body, consisting of the deactivation and excretion of xenobiotics, and happens mainly in the liver. For example, a group of enzymes involved in xenobiotic metabolism is the hepatic microsomal cytochrome P450 system.

Table 1. Examples of *In Vitro* Models Used in Toxicogenomics

Model	Organ	Species	Advantages	Disadvantages	Refs.
Tissue slices	Liver	Rat/mouse/human	• Liver structure is maintained with all cell types • Good correlation with *in vivo* regarding **xenobiotic metabolism** and zone-specific cytochrome activity • **Phase II enzymes**, gluconeogenesis, and albumin production could be retained with 20–96 hours	• Necrosis occurs after 48–72 hours • Metabolic enzyme levels decreased after 6 72 hours • Drug metabolism and intrinsic clearance rates are lower than primary hepatocytes	[52]
Primary hepatocytes	Liver/kidney	Rat/mouse/human	• Functional activities could be maintained for 24–72 hours • Ideal for assessing the interspecies and interindividual differences in metabolism • Suitable for enzyme induction and inhibition studies	• Hepatocyte de-differentiation changes function, gene expression, cell morphology • Microenvironment lost • Cytochrome P450 expression decline quickly after 24–48 hours	[53]
Immortalized cell lines (e.g., HepaRG and HepG2)	Liver	Rat/mouse/human	• High proliferative capacity and stable karyotype • Expression level of most liver functions and phase I/II enzymes can be retained in a lower percentage than primary hepatocytes	• Individual donor phenotype can be retained • Limited predictive power of toxicity is retained in population level	[54]
Three-dimensional culture systems	Liver	Rat/mouse/human	• Hepatocyte functions are improved compared with monolayer culture • Cell types are retained and longevity is extended • Good correlation with *in vivo* toxicity • Cell interaction, morphology is more stable	• Limited successful co-culture (mainly with fibroblasts cells) • Standard culture construction protocol is needed • Not fully high throughput	[55]
Embryonic stem cells	Liver/cardio/brain	Rat/mouse/human	• Easily studied with most established omics techniques • Define phenotypes for many organ toxicity • Developmental toxicity	• Lower expression for metabolism-related genes • Not fit for long-term experiments • Bioengineering required	[51]
Induced pluripotent stem cells	Liver/cardio/brain	Rat/mouse/human	• Individual variability can be assessed • **Idiosyncratic toxicity** • Defined phenotype (multiple disease models) • Easily studied with most established omics techniques	• Bioengineering required • Lack of robust and reproducible differentiation protocols • Loss of the functionality of native hepatocytes • Lower expression level of metabolism-related genes	[56]
Organoids	Multiple organs	Rat/mouse/rat and other species	• Limited amounts of starting material required • Can be propagated for a long time • Can be derived from multiple tissues and species • Good preservation of physiological features	• The native microenvironment of derived tissues could not be well maintained • Unable to mimic *in vivo* growth factor • Limited use in modeling inflammatory responses of tissues	[57]

The Choice of Genomic Methods in TGx Research

NGS offers several benefits beyond gene expression profiling, as most TGx studies are more focused than NGS. For example, NGS has been widely used in detecting novel genetic variants, including complex structural variants such as gene fusion, rare mutation, short and long noncoding RNAs (ncRNAs) [17], and circular RNAs [18] (Figure 2). These novel genomic features are expected to advance the role of TGx in a broad range of toxicology studies. Despite the significant decline in the cost of NGS, it has remained economically challenging to screen vast numbers of industrial and environmental chemicals with the current NGS technologies. Thus, researchers still need to carefully weigh the options of choosing one technology over others by considering cost, purpose, and speed.
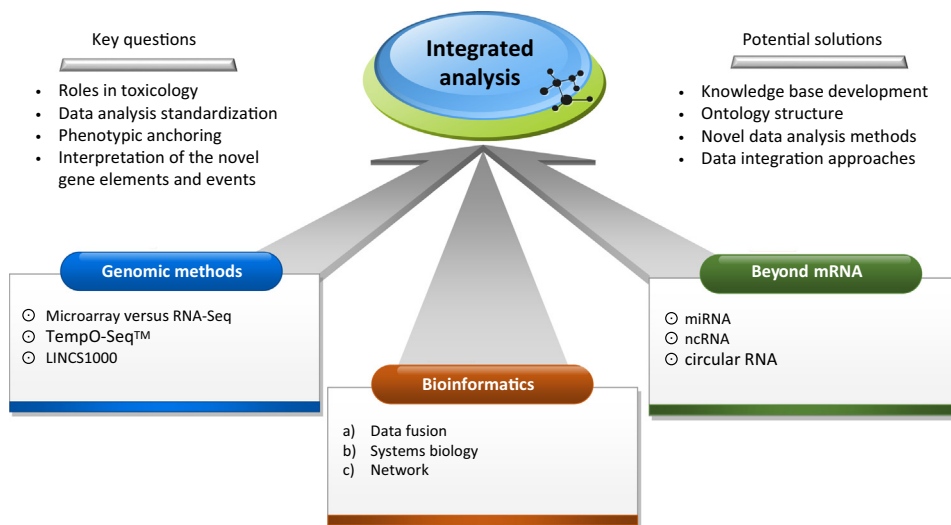
Trends in Pharmacological Sciences

Figure 2. Application of Novel Genomics Technologies in Toxicogenomics (TGx). Some novel genomics technologies such as RNA-Seq and TempO-Seq™ have been well established in the TGx field. The novel genomics features beyond mRNA, including miRNA, noncoding RNA, and circular RNAs, provided extra resolution for identifying and understanding the underlying mechanism of toxicities. Bioinformatics approaches, including data integration, systems biology, and the network, could serve as a bridge to facilitate the utilization of novel genomics technologies in TGx.

Some novel approaches developed and applied in the TGx field are fast, affordable, and applicable to risk assessment. For example, BioSpyder Technologies has developed a novel product for targeted sequencing called TempO-Seq™, a gene expression profiling tool designed to monitor hundreds to thousands of genes at once in high throughput at an extremely lower cost per sample. Furthermore, the technology is very sample friendly, accepting diverse samples, such as formalin-fixed paraffin-embedded tissues and single cells from different species, with minimal sample input (i.e., 10 pg) [19]. Another example is the National Institute of Health Library of Integrated Network-Based Cellular Signatures (LINCS) 1000 project [20], which aims to provide a cost-effective panel-based method profiling 1000 landmark genes that can recapitulate the biology that otherwise needs to be assessed at the whole-genome scale [21]. Based on a similar concept, Tox21 developed the S1500 gene panel, which is scheduled to screen approximately 10 000 chemicals, including most drugs [22].

Beyond mRNA

Advances in genomics and bioengineering technology have provided unprecedented opportunities to identify different genomic features other than mRNAs [23]. These novel components have been demonstrated to have diverse roles in biological processes and have expanded our understanding of the underlying mechanism of toxicity (Figure 2). To advance the utility of these novel gene elements in regulatory toxicology, the following should be addressed: (i) determine the roles of novel genomic features in toxicology and the extent to which they are fit-for-purpose in assessing toxicity; (ii) develop state-of-art data analysis strategies for generating robust and reproducible measures of these new features; and (iii) integrate various molecular events with associated phenotypic outcomes to decipher the gene regulatory complex essential to the expression of toxicity.

Efforts have been made to link the different genetic elements to phenotypes in TGx field [24]. For example, miRNAs are small ncRNAs that have been demonstrated as one of the most stable

RNA types and are conserved across many species. The expression of miRNAs tends to be cell-type specific, which allows the development of biofluid-based miRNA detection assays that work as a 'liquid biopsy' to carry out fast and noninvasive early detection of toxicity [25]. Besides miRNAs, other ncRNAs have also been identified and linked to many biological processes through diverse mechanisms [26]. Furthermore, some integrative approaches, such as data fusion, systems biology, and network analysis have been developed to incorporate multiple genetic elements for risk assessment [27]. For example, Nan *et al.* [28] uncovered the regulatory interactions among ncRNA, circular RNA, and miRNA for lead-induced neurotoxicity and further verified these results with dual luciferase reporter gene assay.

Reproducibility

Reproducibility is the most important hallmark of science [29]. Concerns have been raised on the reproducibility of genomics technology due to the complex nature and aspects regarding cell types, genomics platforms, data analysis methodologies, and an inherent tendency to variability both within and between different laboratories involved in testing and validation (Figure 3). Therefore, a critical evaluation must be carried out on the reproducibility of TGx signatures and their transferability in the biological context [30] (Box 1).

The typical TGx study is designed to determine differentially expressed genes (DEGs) by comparing the treated condition with a control, followed by biological interpretation (e.g., pathways and gene functions) to understand the underlying toxicity mechanism of the treatment. DEGs are extremely sensitive to study design, including the animal strain and batch, cell culture methodologies, treatment protocol, and measured endpoints [12]. This highlights that
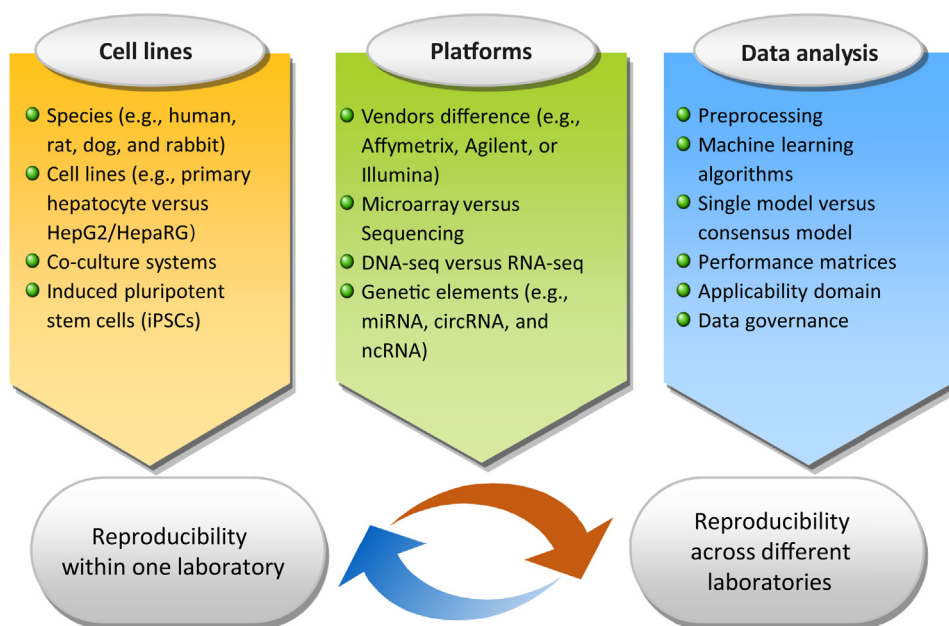
Trends in Pharmacological Sciences

Figure 3. Key Challenges of Reproducibility in Toxicogenomics. The reproducibility of toxicogenomics include both biological and technical sides. The data are generated from different cell lines in various labs with different technologies. The repeatability and reproducibility necessities have different influential factors including cell types, genomics platforms, and data analysis. cirRNA, Circular RNA; ncRNA, noncoding RNA.

Box 1. Key Considerations of TGx in a Regulatory Setting

Reproducibility: best practice in data analysis is required to ensure the reproducibility of results in TGx. To fill the gap, Consortia such as MicroArray/Sequencing Quality Control (MAQC and SEQC) are driving the reproducibility of transcriptional measurement. Some best-practices for TGx data analysis have been suggested: (i) the differentially expressed genes (DEGs) calculated by fold change (1.5) with a *P* value cutoff (0.05) are suggested for use in downstream pathway or gene ontology analysis; (ii) the model performance mainly depended on the complexity of endpoints and team proficiency rather than machine learning algorithms; and (iii) the cross-platform concordance regarding DEGs was correlated with treatment effect, the complexity of biological mechanisms, and transcript abundance. The RNA-Seq outperformed microarray due to low-abundance transcripts, which means the similar performance could be generated from both RNA-Seq and microarray platform when limiting the DEGs in the high expression level [32].

Adverse outcome pathways (AOPs): AOPs have been depicted as biological dominos integrating molecules, gene activities, and causal adverse outcomes in sequence. TGx is an important resource to seamlessly link the information among the AOP components and facilitate assembling knowledge. TGx has been adopted for developing organ toxicity-related AOP, such as neurotoxicity [58] and fatty liver diseases [45]. Furthermore, the Organization for Economic Cooperation and Development established AOP development under the oversight of its Extended Advisory Group on Molecular Screening and Toxicogenomics (EAGMST), which aims to develop guidance detailing internationally accepted approaches for describing and documenting AOPs for the potential regulatory application.

Read-Across: read-across assumes that two 'similar' compounds are likely to share a similar toxicity profile. Currently, the main emphasis is on structural similarity. However, diverse biological data for compounds of interest have been generated and become available. These biological data, including TGx data, could be utilized to profile the biological fingerprint of a chemical. If the two chemicals have similar biological profiles, they have a similar adverse outcome. Biological-based read-across could be a complement to structure-based read-across. Some case reports of biological-based read-across have been described for various complex toxicity endpoints, such as estrogenic endocrine disruption [47].

---

the recording of methods or protocols in TGx is essential for a result to be reproduced by others. Therefore, the community has been discussing the value of incorporating good laboratory practices (GLPs) into TGx. Some initial efforts, such as the Organization for Economic Co-operation and Development, have begun to develop a GLP standard for the conduct of TGx studies in screening compounds for toxicity detection. Even with a standardized and GLP-like protocol, DEGs could still vary significantly with the choice of data processing and statistical methods. This so-called 'computational reproducibility' issue poses a specific and significant challenge to TGx and to the genomics field in general. It is unlikely that there exists a single analysis method that is universally correct for every TGx study and every gene measured in the study. However, a general rule can be followed, as suggested by the FDA-led MicroArray Quality Control (MAQC) consortium, to determine a reproducible DEG profile: ranking the genes by fold-change by comparing the treated with the control condition, followed by removing these genes with poor statistics (measured by a nonstringent *P* value) [31]. The MAQC consortium further demonstrated that, with this simple rule, a reproducible DEG profile was readily achieved across different microarray platforms for the example samples tested [32].

TGx, just like genomics applied in other fields, involves substantial computational approaches to analyze data and visualize/present the results. TGx can learn from other fields, such as the **FAIRsharing community**, to develop a strategy to reproduce the key figures and results from published TGx studies [33]. Several journals, such as *Scientific Data* and *GigaScience,* aim to describe scientifically valuable datasets and research that advances the sharing and reuse of scientific data. Moreover, technology such as blockchain, that aims to create incorruptible data trails and securely record decisions, may play an important role in improving reproducibility in TGx [34]. In addition, the scientific community has pushed the 'open science' concept to enhance data transparency and reproducible science. Some public data and code

repositories, such as Github, have been developed, facilitating data and code share, and communication among researchers.

Machine Learning and Predictive Models

Machine learning has been widely used in the genomics and genetics [35], and it has also enabled TGx to place toxicology on a more predictive footing. Artificial intelligence (AI), including machine learning, has been evaluated in the regulatory setting to facilitate digital health criteria establishment [33]. Such a rapid advancement from research to the regulatory application for AI has been largely encouraged by the success of deep learning in a broad range of biomedical research [36]. As more and more TGx data is generated and becomes available in the near future, AI, particularly deep learning, will play an increasing role in toxicology, specifically for TGx [37]. However, the nature of toxicology studies and the inherent complexity in a TGx study design poses many challenges in developing a meaningful and reliable TGx-based predictive model [38] (Box 2). Firstly, the sample size is usually small in TGx. Hence, it is difficult to assess whether a predictive model developed based on a limited number of samples could well represent a broad space of measured endpoints. Secondly, the interpretation of such a predictive model is difficult. A lot of models have good prediction performance but the features (e.g., gene signatures) employed in the models are hard to link to the existing knowledge of toxicity endpoints. Thirdly, it is difficult to take the entire complexity of a TGx experiment into consideration in predictive model development. For example, many TGx experiments were carried out with multiple times and doses. Therefore, not all of the machine learning methodologies are capable of incorporating these important toxicology features into the predictive modeling framework. Nonetheless, several approaches have been explored, and some of them could be important for the future development and application of deep learning in TGx [33].

Transcriptional Profiling and Image Analysis

One of the most important or appreciated applications of deep learning is for image recognition and classification. Deep convolutional neural networks (CNN) is one of the most successful

---

Box 2. Machine Learning/AI in Toxicogenomics

AI is when a machine is trained to think and act like a human. Machine learning is a field of AI that uses statistical techniques to give computer systems the ability to 'learn' from data, without being explicitly programmed. Deep learning goes beyond machine learning and uses algorithms inspired by the structure and function of the brain (so-called artificial neural networks) to reveal hidden relationships between experimental observations and causal biological endpoints. Learning can be supervised (where each example is a pair of an input and desired output) or unsupervised (where the AI system is presented with uncategorized data).

Machine learning has been used since the inception of TGx to assist in data manipulation and interpretation. However, recent advancements in AI and especially deep learning provide new opportunities to develop biomarkers with enhanced performance and predictive capability. For example, AI approaches could assist with challenges such as overcoming small sample sizes in TGx studies, integrating multiple gene profiles from different cell lines, and in transferring transcriptional profiles to images via the biological relationship between genes. Consequently, deep learning combined with TGx offers new ways to predict complex toxicological endpoints, such as nongenotoxic carcinogenicity, alongside conventional approaches, such as early biomarker research and mechanistic analyses.

A key consideration in using predictive TGx models in the regulatory setting is to define context of use both in terms of the scope of application and the specific roles that TGx models could play in regulatory decision making. Alongside this scope, there are questions to be addressed, such as (i) Biological plausibility: is the gene signature in the model consistent with existing knowledge? (ii) Biological reliability: does the gene signature cross-validate with other methods such as real-time PCR? (iii) Statistical performance: which statistical criterion best defines the reliability of the model? (iv) Reproducibility: what degree of reproducibility is required?

methodologies, since it can achieve high accuracy by emulating the behavior of optic nerves in living creatures. CNN uses mathematical filters to correlate image pixel values with neighboring pixel values to classify and interpret images [39]. A similar concept can be applied to TGx by treating a transcriptional profile in the same way as an image without information loss. Specifically, analog proximity to measure gene–gene relationship can be established to transform the 1D gene expression profiles to 2D images. Some initial attempts have been reported by using pathways, gene ontology structure with distance measure strategies, to reorganize transcriptional profiles to images for deep CNN. For example, Ma and Zhang [40] proposed an **OmicsMapNet** strategy to rearrange 1D transcriptional profiling to the 2D image for deep CNN model development, where the transcriptional profiling was mapped to the **KEGG** pathway and then turned into an image with a **treemap approach**.

Overcoming the Small Sample Size in TGx Studies through Transfer Learning

The central concept of transfer learning is to assume that the knowledge learned from one dataset with abundant information can be transferred to learn from a new dataset with limited information. This concept is specifically relevant in TGx studies where the sample size is always limited, posing a challenge to making reliable conclusions. Transfer learning is widely applied in image analysis [41] and automatic speech recognition [42], but has not been investigated in TGx. Many large genomics datasets such as GEO, Open TG-GATEs, and DrugMatrix have been published in the past decades and can be used as transfer learning to improve the predictive ability of TGx models with small data sets. For example, Chen *et al*. [43] developed a multitask multilayer feedforward neural network to inference the gene expression by using LINCS 1000, Genotype-Tissue Expression (GTEx) data and 1000 Genomes expression data, yielding a lower error rate than state-of-art machine learning algorithms [43].

Unsupervised Learning

Unsupervised learning such as principle component analysis (PCA) is focused on the discovery of natural groupings of subjects (or samples) based on the data itself (e.g., gene expression data from microarrays). This method is of particular relevance in TGx applications since it is always a challenge to define a phenotypic outcome and its association with gene expression changes. Therefore, unsupervised learning like PCA has been extensively used in TGx, particularly when dealing with a TGx experiment involving multiple doses and time points. For example, PCA was employed to choose the most representative time/dose condition in TG-GATEs data [44]. Unsupervised learning in a deep learning framework has brought many opportunities for enhanced knowledge discovery in TGx. One interesting deep learning framework has been proposed to aggregate the genomics data from different experimental conditions for information extraction. This approach has been successfully applied to compare and cluster genomics data of *Pseudomonas aeruginosa* from different experiments conducted in different media conditions [36]. Some TGx related questions that could potentially benefit from this approach include: (i) *in vitro* to *in vivo* extrapolation (**IVIVE**), (ii) utilization of big genomics data from immortal cells to enhance toxicity prediction, and (iii) elimination of the cell culture discrepancy for integrative model development.

The Role of TGx in Developing AOPs

Risk assessment of chemical exposure is mostly based on understanding the underlying mechanism of exposure and its relevance to human health. To facilitate this line of research, the concept of AOPs has been established; this provides a conceptual framework to link chemical exposure to adverse outcome or key event responses at the cellular, organ, or organism levels via molecular initiating events and perturbation pathways. To date, a total of

234 AOPs have been reported that correspond to 1764 key events. AOPs have been actively studied to address different toxicity endpoints, such as drug-induced liver injury, liver carcinogenicity, and neurotoxicity. Curation of an AOP concept requires tremendous knowledge and is time-consuming. Accumulative TGx data profiles could accelerate the AOP concept generation for community effort to verify the hypotheses (Box 1). These resources can be integrated with the large publicly available TGx datasets, such as Open TG-GATEs [11] and DrugMatrix, to develop *in silico* AOP concepts. For example, Bell *et al*. [45] incorporated Open TG-GATEs data sets, ToxCast *in vitro* assays, and rat protein–protein interactions in the Reactome database to identify computationally predicted AOP (cpAOP) scaffolds. The cpAOP can serve as a starting point for domain experts to accelerate the AOP development.

Read-Across with TGx

Read-across assumes that two 'similar' compounds are likely to share a similar toxicity profile [46]. Read-across has been actively promoted in risk assessment of chemical exposure by offering a route to 'refine, reduce and replace' animal use in **Registration, Evaluation, Authorization and Restriction of Chemicals** (**REACH**) [47]. For example, the Research Institute for Fragrance Materials has adopted read-across for assessing the risks of 80% of fragrance ingredients. Conventional read-across has mainly relied on chemical structure similarity [47], which has met with challenges since the chemistry space usually cannot represent the complexity of a biological process. A lot of false positive results were generated by chemical structure-based read-across, which lacks consensus with external evidence to substantiate a read-across for a given endpoint such as reproductive and development toxicity [48].

With the help of 'big data' and omics profiles, the performance of read-across can be improved. Accumulative TGx data sets can be employed to establish the similarity for projecting the shared toxicities profiles (Box 1). Some TGx applications, such as DrugMatrix and LINCS 1000, allows users to compare the measured compound with a huge number of tested compounds at the transcriptomic levels. Furthermore, a consensus similarity approach (where similarities are aggregated from different biological data profiles) could further enhance the reliability and reproducibility of read-across. For example, Zhu *et al*. [46] conducted a read-across using biological data, including TGx data, for assessing acute toxicity within the REACH program and extrapolation between *in vitro* and short-term *in vivo* repeated dose studies with stem cells.

Concluding Remarks and Future Perspectives

The rapid growth of genomics techniques has dramatically extended the impact and reach of TGx. However, challenges in experimental design, statistical interpretation, and reproducibility need to be addressed before TGx can realize its full potential (see Outstanding Questions). There are growing calls to drive more reproducibility in sciences, avoiding some wildly optimistic and misleading conclusions [29]. The Massive Analysis and QC Society, derived from the FDA-led consortium activities, aims to communicate, promote, and advance reproducible science principles and quality control for analysis of massive data generated from the existing and emerging technologies in solving biological, health, and medical problems [30]. Besides technical reproducibility, the dimension of reproducible sciences in TGx also includes understanding the TGx findings and how the TGx results correlate with toxicity in humans. It is expected that more benchmark TGx data will be generated under different platforms, cell types, and cultures to facilitate an evaluation and robust conclusion generation.

Accumulative TGx data flooding into the community has accelerated the molecular understanding of toxicity and enhanced predictive capability. Furthermore, some emerging

Outstanding Questions

How can we establish the standards in regulatory oversight of *in vitro* toxicogenomics (TGx) for next-generation risk assessment?

How can we systematically integrate the multiple genomics features to enable an augmented elucidation of complex toxicity outcome?

How can we enhance the reproducibility of TGx results towards more robust and reliable results for regulatory application?

How do we choose the correct 'fit-for-purpose' deep learning approaches to improve predictive power in TGx?

technologies, including organ-on-a-chip [49], 3D-printed tissues [50], and nanoparticle techniques [51], are used increasingly in toxicology. Accordingly, novel methodologies to model, integrate, and mine these valuable resources are urgently needed. Some systems biology and network approaches are promising tools for TGx data profiling and uncovering the complex regulation among different gene elements. Furthermore, novel machine learning algorithms, such as deep learning, may produce more accurate and robust prediction results from TGx studies. Importantly, efforts are needed to consolidate and link the accumulative TGx findings to create information-rich tools and networks for enabling data standardization, consistency, and maintenance.

TGx has gone through a series of peaks since its inception. Now is the right time to rethink and reevaluate TGx technology as a driver to move away from *in vivo* data to new approaches in regulatory decision making. In the face of strategic shifts in the risk assessment paradigm, TGx could play an important role and potentially be widely applied in support of regulatory strategy. Significantly, despite the complexity and challenging nature of TGx, the discipline has exceeded expectations of utility in prediction. Here, we have summarized the challenges and opportunities presented by TGx from a data science perspective to stimulate community efforts for further evaluation and to better position TGx in the regulatory setting.

Disclaimer Statement

The views presented in this article do not necessarily reflect current or future opinion or policy of the US Food and Drug Administration and National Institute of Health. Any mention of commercial products is for clarification and not intended as an endorsement. Dr Ruth Roberts is co-founder and co-director of Apconix, an integrated toxicology and ion channel company that provides expert advice on nonclinical aspects of drug discovery and drug development to academia, industry, and not-for-profit organizations.

Resources

[i]http://www.seurat-1.eu/
[ii]https://www.ncbi.nlm.nih.gov/geo/
[iii]http://toxico.nibio.go.jp/english/index.html
[iv]https://ntp.niehs.nih.gov/results/toxfx/index.html
[v]https://biospyder.com/technology/
[vi]http://www.oecd.org/chemicalsafety/testing/toxicogenomics.htm
[vii]https://desktop.github.com/
[viii]https://commonfund.nih.gov/gtex
[ix]http://www.internationalgenome.org/
[x]https://aopkb.oecd.org/
[xi]http://fragrancematerialsafetyresource.elsevier.com/

References

1. Patlewicz, G. *et al.* (2013) Use of category approaches, read-across and (Q)SAR: general considerations. *Regul. Toxicol. Pharmacol.* 67, 1–12

2. Barroso, J. *et al.* (2017) Cosmetics Europe compilation of historical serious eye damage/eye irritation *in vivo* data analysed by drivers of classification to support the selection of chemicals for development and evaluation of alternative methods/strategies: the Draize eye test Reference Database (DRD). *Arch. Toxicol.* 91, 521–547

3. Hamburg, M.A. (2011) Advancing regulatory science. *Science* 331, 987

4. Attene-Ramos, M.S. *et al.* (2013) The Tox21 robotic platform for the assessment of environmental chemicals – from vision to reality. *Drug Discov. Today* 18, 716–723

5. Dix, D.J. *et al.* (2007) The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicol. Sci.* 95, 5–12

6. Sewell, F. *et al.* (2018) The future trajectory of adverse outcome pathways: a commentary. *Arch. Toxicol.* 92, 1657–1661

7. Nel, A. *et al.* (2013) Nanomaterial toxicity testing in the 21 st century: use of a predictive toxicological approach and high-throughput screening. *Acc. Chem. Res.* 46, 607–621

8. Boess, F. *et al.* (2003) Gene expression in two hepatic cell lines, cultured primary hepatocytes, and liver slices compared to the *in vivo* liver gene expression in rats: possible implications for toxicogenomics use of *in vitro* systems. *Toxicol. Sci.* 73, 386–402

9. Sutherland, J.J. *et al.* (2016) Assessing concordance of drug-induced transcriptional response in rodent liver and cultured hepatocytes. *PLoS Comput. Biol.* 12, e1004847

10. Liu, Z. *et al.* (2018) Transcriptional responses reveal similarities between preclinical rat liver testing systems. *Front. Genet.* 9, 74

11. Igarashi, Y. *et al.* (2015) Open TG-GATEs: a large-scale toxicogenomics database. *Nucleic Acids Res.* 43, D921–D927

CellPress
REVIEWS

12. SEQC/MAQC-III Consortium (2014) A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat. Biotechnol.* 32, 903

13. Lauschke, V.M. *et al.* (2016) Novel 3D culture systems for studies of human liver function and assessments of the hepatotoxicity of drugs and drug candidates. *Chem. Res. Toxicol.* 29, 1936–1955

14. Rostami-Hodjegan, A. (2018) Reverse translation in PBPK and QSP: going backwards in order to go forward with confidence. *Clin. Pharmacol. Ther.* 103, 224–232

15. Merrick, B.A. *et al.* (2015) Intersection of toxicogenomics and high throughput screening in the Tox21 program: an NIEHS perspective. *Int. J. Biotechnol.* 14, 7–27

16. Shendure, J. *et al.* (2017) DNA sequencing at 40: past, present and future. *Nature* 550, 345

17. Dempsey, J.L. and Cui, J.Y. (2017) Long non-coding RNAs: a novel paradigm for toxicology. *Toxicol. Sci.* 155, 3–21

18. Holdt, L.M. *et al.* (2018) Circular RNAs as therapeutic agents and targets. *Front. Physiol.* 9, 1262

19. Yeakley, J.M. *et al.* (2017) A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling. *PLoS One* 12, e0178302

20. Duan, Q. *et al.* (2014) LINCS Canvas Browser: interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures. *Nucleic Acids Res.* 42, W449–W460

21. Keenan, A.B. *et al.* (2018) The Library of Integrated Network-Based Cellular Signatures NIH Program: system-level cataloging of human cells response to perturbations. *Cell Syst.* 6, 13–24

22. Mav, D. *et al.* (2018) A hybrid gene selection approach to create the S1500+ targeted gene sets for use in high-throughput transcriptomics. *PLoS One* 13, e0191105

23. Wilkes, M.C. *et al.* (2017) Beyond mRNA: the role of non-coding RNAs in normal and aberrant hematopoiesis. *Mol. Genet. Metab.* 122, 28–38

24. Sutherland, J.J. *et al.* (2017) Toxicogenomic module associations with pathogenesis: a network-based approach to understanding drug toxicity. *Pharmacogenomics J.* 18, 377–390

25. Harrill, A.H. *et al.* (2016) MicroRNA biomarkers of toxicity in biological matrices. *Toxicol. Sci.* 152, 264–272

26. Gupta, S.K. *et al.* (2017) Quaking inhibits doxorubicin-mediated cardiotoxicity through regulation of cardiac circular RNA expression. *Circ. Res.* 122, 246–254

27. Caiment, F. *et al.* (2015) High-throughput data integration of RNA–miRNA–circRNA reveals novel insights into mechanisms of benzoapyrene-induced carcinogenicity. *Nucl. Acids Res.* 43, 2525–2534

28. Nan, A. *et al.* (2017) A novel regulatory network among LncRpa, CircRar1, MiR-671 and apoptotic genes promotes lead-induced neuronal cell apoptosis. *Arch. Toxicol.* 91, 1671–1684

29. Munafò, M.R. *et al.* (2017) A manifesto for reproducible science. *Nat. Hum. Behav.* 1, 0021

30. Shi, L. *et al.* (2017) The international MAQC Society launches to enhance reproducibility of high-throughput technologies. *Nat. Biotechnol.* 35, 1127

31. MAQC Consortium *et al.* (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* 24, 1151

32. Wang, C. *et al.* (2014) The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nat. Biotechnol.* 32, 926

33. Sansone, S.-A. *et al.* (2018) FAIRsharing: working with and for the community to describe and link data standards, repositories and policies. *bioRxiv* Published online January 17, 2018. http://dx.doi.org/10.1101/245183

34. Chapron, G. (2017) The environment needs cryptogovernance. *Nature* 545, 403–405

35. Libbrecht, M.W. and Noble, W.S. (2015) Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* 16, 321

36. Tan, J. *et al.* (2017) Unsupervised extraction of stable expression signatures from public compendia with eADAGE. *bioRxiv* Published online April 10, 2017. http://dx.doi.org/10.1101/078659

37. Mayr, A. *et al.* (2016) DeepTox: toxicity prediction using deep learning. *Front. Environ. Sci.* 3, 80

38. Qin, C. *et al.* (2016) Toxicogenomics in drug development: a match made in heaven? *Expert Opin. Drug Metab. Toxicol.* 12, 847–849

39. Litjens, G. *et al.* (2017) A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88

40. Ma, S. and Zhang, Z. (2018) OmicsMapNet: transforming omics data to take advantage of Deep Convolutional Neural Network for discovery. *arXiv* 1804.05283

41. Simonyan, K. and Zisserman, A. (2014) Very deep convolutional networks for large-scale image recognition. *arXiv* 1409.1556

42. Dahl, G.E. *et al.* (2012) Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Trans. Audio Speech Lang. Process.* 20, 30–42

43. Chen, Y. *et al.* (2016) Gene expression inference with deep learning. *Bioinformatics* 32, 1832–1839

44. Liu, Z.C. *et al.* (2017) *In vitro* to *in vivo* extrapolation for drug-induced liver injury using a pair ranking method. *ALTEX* 34, 399–408

45. Bell, C.C. *et al.* (2016) Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease. *Sci. Rep.* 6, 25187

46. Zhu, H. *et al.* (2016) Supporting read-across using biological data. *ALTEX* 33, 167–182

47. Ball, N. *et al.* (2016) Toward good read-across practice (GRAP) guidance. *ALTEX* 33, 149–166

48. Leist, M. *et al.* (2014) Consensus report on the future of animal-free systemic toxicity testing. *ALTEX* 31, 341–356

49. Huh, D. *et al.* (2012) A human disease model of drug toxicity–induced pulmonary edema in a lung-on-a-chip microdevice. *Sci. Transl. Med.* 4, 159ra147

50. Vanderburgh, J. *et al.* (2017) 3D printing of tissue engineered constructs for *in vitro* modeling of disease progression and drug screening. *Ann. Biomed. Eng.* 45, 164–179

51. Gao, X. *et al.* (2017) Toxicity of nano- and ionic silver to embryonic stem cells: a comparative toxicogenomic study. *J. Nanobiotechnology* 15, 31

52. Vatakuti, S. *et al.* (2016) Classification of cholestatic and necrotic hepatotoxicants using transcriptomics on human precision-cut liver slices. *Chem. Res. Toxicol.* 29, 342–351

53. Boess, F. *et al.* (2017) Use of early phenotypic *in vivo* markers to assess human relevance of an unusual rodent non-genotoxic carcinogen in vitro. *Toxicology* 379, 48–61

54. Krivoshiev, B.V. *et al.* (2018) Toxicogenomics of the flame retardant tris (2-butoxyethyl) phosphate in HepG2 cells using RNA-seq. *Toxicol. In Vitro* 46, 178–188

55. Bell, C.C. *et al.* (2017) Transcriptional, functional, and mechanistic comparisons of stem cell–derived hepatocytes, HepaRG cells, and three-dimensional human hepatocyte spheroids as predictive *in vitro* systems for drug-induced liver injury. *Drug Metab. Dispos.* 45, 419–429

56. Choudhury, Y. *et al.* (2017) Patient-specific hepatocyte-like cells derived from induced pluripotent stem cells model pazopanib-mediated hepatotoxicity. *Sci. Rep.* 7, 41238

57. Fatehullah, A. *et al.* (2016) Organoids as an *in vitro* model of human development and disease. *Nat. Cell Biol.* 18, 246

58. Bal-Price, A. and Meek, M.E. (2017) Adverse outcome pathways: application to enhance mechanistic understanding of neurotoxicity. *Pharmacol. Ther.* 179, 84–95